

ADDING CMS TIER 3 COMPUTING TO AN ALREADY EXISTING CLUSTER AT TEXAS A&M UNIVERSITY

Guy Almes¹, Daniel Cruz¹, Jacob Hill²,
Steve Johnson¹, Michael Mason^{1 3},
Vaikunth Thukral¹, David Toback¹, Joel
Walker²

1. Texas A&M University
2. Sam Houston State University
3. Los Alamos National Laboratory

Outline

- Overview of CMS grid computing
- Tier3 site at Texas A&M
 - Tier3 site on an already existing cluster
- Advantages & Disadvantages to installing a Tier3 on an existing cluster
- Performance
- Custom monitoring tools for Tier3 sites
 - Why we need custom monitoring
- Conclusions

CMS Grid Computing

- Computers → Clusters → Grid
- CMS Grid
 - Tiered (layered) structure (by functionality)
- How CMS Grid handles data:
 - PhEDEx
 - Data around world; bring over tens of TB's per month
 - CRAB
 - How jobs are submitted to the grid

Organization of the CMS Tiers

- Why particular structure? Better allocation & handling of resources
- Tier 0: CERN
- Tier 1: A few National Labs
- Tier 2: Bigger University Installations for national use (shared resources)
- Tier 3: Local use (our type of center; balance between meeting our users' needs and helping community)

Installing a Tier3 on an already existing, shared cluster

- Worth noting: most people just create their own cluster
- When the time came for the A&M CMS Group to add Grid Computing capabilities, we decided to join an already existing cluster
- Brazos is a computing cluster at the university
 - Designed for high throughput (as opposed to high performance)
- Mostly used by stakeholders in the College of Science

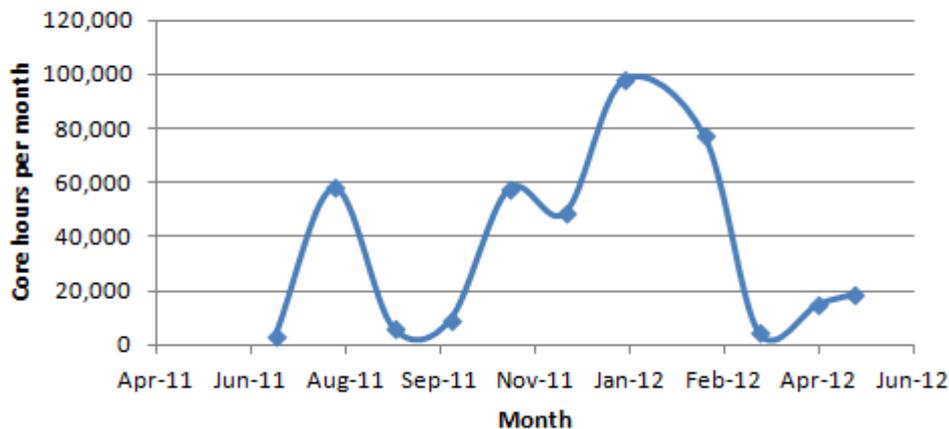
Advantages & Disadvantages

- Advantages:
 - Already established support and System administration; we have 256 *dedicated* priority cores, but we can run on all 2565 cores if needed
- Disadvantages
 - Dealing with University firewall issues; operating systems for various users; system administrators who want us to change how we run on the Grid; we want to let users from around the globe run if needed, but shared clusters often have rules about who can run

Going Live: Performance

- A&M Tier3 site came online on December 2010
- Spent first few months of 2011 dealing with testing woes
- Really took off around August 2011

HEPX Usage



A plot of the core-hours per month run by our HEPX group (April 2011 – May 2012); dips due to many factors: problems with cluster, user inactivity, etc.

Monitoring

- CMS does a *great* job of monitoring for Tier 0's, 1's, and 2's, as well as jobs and data transfers
 - Tier3 monitoring is generally unsupported by CMS
- Tier3's mostly supported by user community
- Decided to put in place monitoring specific for our Tier3 site (and transferable to other Tier3's)
 - Data transfers and load tests working?
 - Cluster up and running?
 - Test jobs working quickly? (custom setup by TAMU)
 - Users jobs working?
- Standard CMS monitoring is on lots of pages → we combined everything you need in one main webpage with all the info
 - Most important: now we have automated checks on the status, send emails if anything goes wrong

Monitoring

- Need monitoring to handle *many* types of debugging
- Example: User has generic complaint: “my jobs aren’t working”
- How do we figure out where the problem is?
 - User problem? Permissions problem? CMS software problem? Incompatibility between software & cluster?
 - Right software installed? Is cluster up and running? Is the cluster connecting to the Grid?
- Did someone, somewhere in the world, change anything? Lots of things need to work just for a job to run to completion

Monitoring

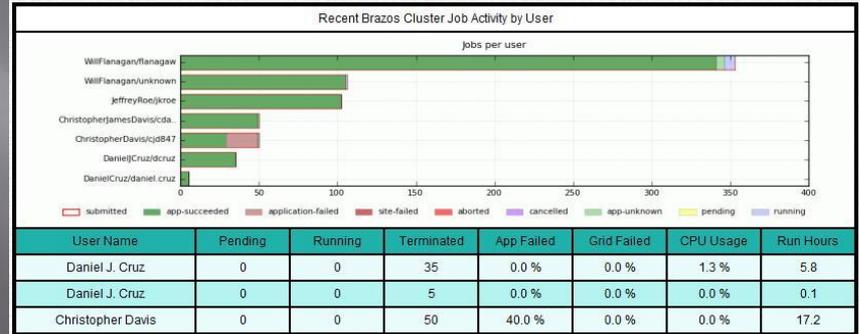
- These are all INTERFACE problems:
 - System administrators can't do it alone → don't know CMS software
 - Physicists can't do it alone → don't know cluster administration or grid infrastructure
- Goal is to provide LOTS of standardized checks, see if we can quickly locate the general source of the problem
- This has led to our custom monitoring website
 - Fully functional website is now in use → <http://collider.physics.tamu.edu/tier3/mon/>

Monitoring

- This has enabled us to successfully install a Tier 3 site on an existing cluster
- Because these tools incorporate our experience, they prove useful to other T3's
- Working on this project at the moment → currently installing in Texas Tech

TORQUE Queue Job Status (2012-06-08 05:25 UTC)								
User Name	Queue	Run Status	Processors	Memory	CPU Hours	Run Hours	Queue Hours	
→ Will Flanagan	HEPX	8 / 8	8	6.0 GiB	124.9	128.7	0.1	
Totals	HEPX	8 / 8	8	6.0 GiB	124.9	128.7	0.1	
↑ Click to Expand or Collapse Table			Select Sorting By →		User	Queue	Status	

Condor Queue Job Status (2012-06-08 05:25 UTC)								
User Name	Universe	Run Status	Processors	Memory	CPU Hours	Run Hours	Queue Hours	
Totals	-	- / -	-	-	-	-	-	-
No Jobs Queued			Select Sorting By →		User	Universe	Status	



Grid Client	CRAB Analysis Test Suite (CATS) Completed Job Status			
Output Host →	Local: Brazos Cluster		Remote: Fermi National Laboratory	
Output Size →	Small	Large	Small	Large
gLite	Pass:10 Fail:0 Other:0 2012-06-07 10:03 UTC	Pass:10 Fail:0 Other:0 2012-06-06 16:03 UTC	Pass:10 Fail:0 Other:0 2012-06-07 22:03 UTC	Pass:10 Fail:0 Other:0 2012-06-08 04:04 UTC
CondorG	Pass:0 Fail:0 Other:10 2012-06-07 09:02 UTC	Pass:0 Fail:0 Other:10 2012-06-07 13:02 UTC	Pass:10 Fail:0 Other:0 2012-06-07 19:03 UTC	Pass:0 Fail:0 Other:10 2012-06-08 01:03 UTC
PBS	Pass:10 Fail:0 Other:0 2012-06-04 20:42 UTC			
↑ Test Results Link to Job Details ↑				

Conclusions

- Successfully created a Tier3 site on an existing cluster at Texas A&M. Fully functional and all the advantages of a big, well supported system
- Have developed a powerful new monitoring system that alerts problems with email in real time, and provides an easy interface of data for debugging
- Our experiences in bringing up a Tier3 on an existing cluster provide an excellent model for other institutions, and our monitoring tools are readily ported to other institutions (in progress)

The End

Backups

Abstract

- "In this talk, we will present a brief overview of the CMS Tier3 site at Texas A&M University. It is largely unique in that we added resources to an already-existing cluster to create our site as opposed to creating a stand-alone system. We will comment on some of the particulars of our site, the advantages and disadvantages of this choice, as well as how it has performed. We will also discuss some powerful new custom monitoring tools we've developed to optimize our cluster performance."